ISSN 1995-0802, Lobachevskii Journal of Mathematics, 2018, Vol. 39, No. 3, pp. 413–423. © Pleiades Publishing, Ltd., 2018.

Asymptotic Expansion of D-Risks for Hypothesis Testing in Bernoulli Scheme

A. A. Zaikin*

(Submitted by A. I. Volodin)

Kazan Federal University, Institute of Computational Mathematics and Information Technologies, Department of Mathematical Statistics, ul. Kremlevskaya 18, Kazan, 420008 Russia

Received August 28, 2017

Abstract—Currently there is not a single asymptotic expansion of d-risk of any statistical procedure. In this paper we present the asymptotic expansion of d-risk function of the optimal test for Bernoulli scheme. The expansion derivation is based on the Edgeworth series of the sufficient statistic for Bernoulli scheme and therefore it is applicable to any one-parametric exponential model.

DOI: 10.1134/S1995080218030204

Keywords and phrases: *Bayesian analysis, d-risk, asymptotic expansion, Bernoulli distribution.*

1. INTRODUCTION

The problem of asymptotic behaviour of d-risk has been studied in the mathematical literature. These works include [1–5]. All of these works utilise some sort of Bernstein-von Mises theorem, which states asymptotic normality of posterior distribution. Since d-risk is a conditional expectation of posterior distribution with respect to decision function, this approach is natural. Furthermore, most optimal statistical procedures depend on posterior risk or are derived from posterior risk function, so using the posteior asymptotic is being ecouraged.

In regards to asymptotic expansions of d-risks, there are currently no works on this subject. That is due to a compicated nature of conditional expectation, which applies not only to the d-risk, but also to the posterior distribution. Now we want to describe the ways to compute d-risk. And first we need to introduce notation.

Consider i.i.d. variables $\mathbf{X}^{(n)} = (X_1, \dots, X_n)$, $X_i \sim \mathbf{P}_\theta, \theta \in \Theta$, where Θ is some parameter set. Parameter θ is given by a random variable ϑ with a known distribution $\mathbf{G}$. Consider the problem with decision set $\mathcal{D}$ and loss function $l(\theta, d) : \Theta \times \mathcal{D} \mapsto \mathbb{R}^+$. Suppose we have a decision function $\delta_n(\mathbf{X}^{(n)})$ with values in the decision set $\mathcal{D}$. The d-risk function of δ_n is defined as the following expression:

$$\mathcal{R}_{\delta_n}(d) = \mathbf{E}\left\{l(\vartheta, \delta_n) | \delta_n = d\right\}, \quad d \in \mathcal{D}. \tag{1}$$

If we know the form of δ_n, we sometimes can calculate this function directly. This is possible, for example, in the case of $\mathcal{D} \subset \mathbb{R}^m$, so the regular conditional distributions exist. This method will be discussed below, and, ultimately, it is the one applied in this work.

There is also another way to calculate the d-risk. Due to conditional expectation properties, we have

$$\mathcal{R}_{\delta_n}(d) = \mathbf{E}\left\{R_{\delta_n}\left(\mathbf{X}^{(n)}\right) | \delta_n = d\right\}, \quad R_{\delta_n}\left(\mathbf{X}^{(n)}\right) = \mathbf{E}\left\{l(\vartheta, \delta_n) | \mathbf{X}^{(n)}\right\},$$

where $R_{\delta_n}(\mathbf{X}^{(n)})$ is a posterior risk. If posterior risk (or its asymptotic) depends only on δ_n, then the computation is straightforward.

*E-mail: Kaskrin@gmail.com